

Claims

What is claimed is:

5 ⁵⁰ 1. A computer-based method of processing an electronic document generated in accordance with a handwriting system, the method comprising the steps of:

obtaining electronic ink data from the handwriting system, the ink data being associated with the electronic document; and

10 automatically identifying, using at least a portion of the electronic ink data, one or more potential page breaks for possible insertion in the electronic document to maintain a page correspondence between the electronic document and a physical document also generated in accordance with the handwriting system.

2. The method of claim 1, wherein the handwriting system is a personal digital notepad.

3. The method of claim 1, further comprising the step of automatically inserting the one or more identified potential page breaks in the electronic document.

15 4. The method of claim 1, further comprising the step of presenting the one or more identified potential page breaks to a user for approval to automatically insert the one or more identified potential page breaks in the electronic document.

5. The method of claim 1, wherein the step of automatically identifying one or more potential page breaks further comprises the steps of:

20 measuring a spatial difference between consecutive pairs of strokes made in accordance with the handwriting system; and

labeling spatial differences not below a threshold value as possible insertion points.

0052493 072560

6. The method of claim 5, further comprising the steps of:

for each spatial difference not below the threshold value, computing one of the number of strokes and the total arc length associated with strokes that occur between the current possible insertion point and one of the next page breaks in the electronic document and the end of the stroke set of the document; and

assigning one of the number of strokes and the total arc length as a score to be associated with the possible insertion point.

7. The method of claim 6, wherein the assigning step is performed when one of the number of strokes and the total arc length is not below a second threshold value.

8. The method of claim 1, wherein the step of automatically identifying one or more potential page breaks further comprises the steps of:

measuring a temporal difference between consecutive pairs of strokes made in accordance with the handwriting system; and

labeling temporal differences not below a threshold value as possible insertion points.

9. The method of claim 8, further comprising the step of, for each temporal difference not below the threshold value, assigning a score to the corresponding possible insertion point based on a distance from the temporal difference to the threshold value.

10. The method of claim 1, wherein the step of automatically identifying one or more potential page breaks further comprises the step of identifying as a possible insertion point a point before a stroke, made in accordance with the handwriting system, wherein the stroke falls within a constrained region on a page associated with the document and wherein the stroke is not immediately preceded by another stroke in the same constrained region.

11. The method of claim 1, further comprising the step of determining a confidence measure for the potential page break associated with the possible insertion point.

12. The method of claim 1, wherein, in an electronic document that corresponds to a form with fields, the step of automatically identifying one or more potential page breaks further comprises the steps of:

computing a measure of field appropriateness for each stroke made in accordance with the handwriting system to indicate how well a stroke fits within a particular field; and

for N consecutive strokes which do not fit the field of a particular page of the document, identifying a potential page break before these N consecutive strokes.

13. The method of claim 12, wherein the potential page break indicates the page number of the page having a field with which the N consecutive strokes appropriately fit.

14. The method of claim 1, wherein the step of automatically identifying one or more potential page breaks further comprises the steps of:

computing a measure of overlap for each stroke with a previous stroke; and

for N consecutive strokes with a total measure of overlap which is not less than a threshold value, identifying a potential page break before these N consecutive strokes.

15. The method of claim 1, wherein the step of automatically identifying one or more potential page breaks further comprises the steps of:

computing a moving average of spatial positions of strokes, made in accordance with the handwriting system, on a page using a predetermined window width, the computation of the moving average resulting in a spatial position moving average curve;

computing a moving average of a slope associated with the spatial position moving average curve, the computation of the moving average resulting in a slope moving average curve; and

5 identifying one or more negative slopes in the slope moving average curve as potential page breaks.

16. The method of claim 15, wherein negative slopes are identified as potential page breaks based on peak-heights.

17. The method of claim 1, wherein the step of automatically identifying one or more potential page breaks further comprises the steps of:

10 computing a moving average of temporal positions of strokes, made in accordance with the handwriting system, on a page using a predetermined window width, the computation of the moving average resulting in a temporal position moving average curve;

15 computing a moving average of a slope associated with the temporal position moving average curve, the computation of the moving average resulting in a slope moving average curve; and

identifying one or more positive slopes in the slope moving average curve as potential page breaks.

20 18. The method of claim 16, wherein positive slopes are identified as potential page breaks based on peak-heights.

19. The method of claim 1, wherein the step of automatically identifying one or more potential page breaks further comprises the steps of utilizing a learning algorithm.

20. The method of claim 1, wherein the step of automatically identifying one or more potential page breaks further comprises the steps of:

performing two or more scoring procedures, each scoring procedure generating a list whose elements include a possible insertion point and a corresponding score;

5 merging the lists generated by the two or more scoring procedures to form a combined list; and

selecting one or more top scoring possible insertion points as the one or more potential page breaks.

10 21. The method of claim 20, wherein the selecting step further comprises selecting a number of top scoring possible insertion points to match the number of expected pages of the document.

15 22. The method of claim 20, wherein the merging step further comprises merging the lists by one of (i) accepting all possible insertion points except duplicates; (ii) unbiased voting; (iii) biased voting; (iv) using disjunctive normal forms; and (v) using neural networks.

23. The method of claim 1, wherein the step of automatically identifying one or more potential page breaks further comprises the step of identifying a potential page break as a point offset from a possible insertion point determined in accordance with a scoring procedure.

20 24. Apparatus for processing an electronic document generated in accordance with a handwriting system, the apparatus comprising:

at least one processor operative to: (i) obtain electronic ink data from the handwriting system, the ink data being associated with the electronic document; and (ii) automatically identify, using at least a portion of the electronic ink data, one or more

potential page breaks for possible insertion in the electronic document to maintain a page correspondence between the electronic document and a physical document also generated in accordance with the handwriting system; and

a memory, coupled to the at least one processor, which stores the electronic ink data associated with the electronic document.

25. An article of manufacture for processing an electronic document generated in accordance with a handwriting system, comprising a machine readable medium containing one or more programs which when executed implement the steps of:

obtaining electronic ink data from the handwriting system, the ink data being associated with the electronic document; and

automatically identifying, using at least a portion of the electronic ink data, one or more potential page breaks for possible insertion in the electronic document to maintain a page correspondence between the electronic document and a physical document also generated in accordance with the handwriting system.